

# SEQUENCE LISTING

<110> Morsey, Mohamad

<120> NOVEL MELANOCORTIN-4 RECEPTOR SEQUENCES AND  
SCREENING ASSAYS TO IDENTIFY COMPOUNDS USEFUL  
IN REGULATING ANIMAL APPETITE AND METABOLIC RATE

<130> PC10743A

<140> 60/213,909

<141> 2000-06-26

<150> 60/213,909

<151> 2000-06-26

<160> 6

<170> PatentIn Ver. 2.1

<210> 1

<211> 1708

<212> DNA

<213> Feline MC4R Nucleotide Sequence

<400> 1

```
cataaaatca gcagcagcta ctaacactca aagcaatgct tcaggttggg aactaatacc 60
tcagagggcag ctggtgtgaa catgcaaaca cggattcagc tcccagtggc acagcagcca 120
ctaggaaaat tattttgaaa agacctgact gaatgcotca ggctaaagtt aaggtggaag 180
ggaggacaga aaagcaaaga gcagactctt tcaactgaga atgagtattt cagaagccta 240
agattttaca atgaagggtga tcagagccgt tcctgggaga cagtaaaaac tccatttcca 300
gcctgggagc acgtgacatt tactcacaac aggcattgcca atttcagcct canaactttc 360
gggcagacaa aggcgtggag aaaaacactg aggcctacctg acccgagaga tcgaatcaat 420
tcagagggga tctgaatcca ctggtgcagg atgaactcca ctcatcacca tggaatgcac 480
acttctctcc acttctggaa cgcagcacc tacggaccgc acagcaatgc cagtgagtcc 540
cttggaaaag gctactctga tggaggggtgt tatgagcaac tttttgtctc ccttgagggtg 600
tttgtgactc tgggtgtcat cagcttgttg gagaatatcc tgggtgattgt ggcaatagcc 660
aagaacaaaa acctgcattc gcccatgtac tttttcatct gcagcctggc tgtggctgat 720
atgttggtga gcgtgtcaaa cggatccgaa accattgtca tcaccctatt aaacagtaca 780
gatacggacg cgcagagttt caccgtgaat attgataatg tcattgactc ggtgatctgt 840
agctccttgc ttgcatcgat ttgcagcctg ctctcaattg cagtggacag gtactttact 900
atcttttatg ctctccagta ccataacatc atgacggtca ggcgggttgg gatcatcata 960
agttgtatct gggcagcttg cacggtttcg ggcgttttgt tcatcatcta ctacagacagc 1020
agtgtgttca tcatctgcct catcaccatg ttcttcacca tgettggtct catggcctct 1080
ctctatgtcc acatgttctt catggccaga ctgcacatta agagaattgc tgtcctcccg 1140
ggcactggca ccatccgcca aggggccaac atgaagggtg caattaccct gaccatactg 1200
attggggctt ttgttgtctg ctggggcccc ttcttctctc acttaaatatt ctacatctct 1260
tgtccccaga atccttactg tgtgtgcttc atgtctcact ttaacctgta tctcatactg 1320
```

atcatgtgta	attccatcat	cgaccctcta	atztatgcac	tccggagcca	agaactaagg	1380
aaaaccttca	aagagatcat	ctgttgctat	cctctaggcg	gcctctgtga	tttgtctagc	1440
agatactaac	tgtgcagata	gaaacgtgca	taagagactt	cttcattctt	acagaaccgg	1500
aacattgtgc	tttgatgacc	cttttctcct	ctgtgtaagg	catgggttga	gactatctgt	1560
tgtataaatt	taagttcatg	actttttttt	ggaatggaaa	caatgcccag	tctctgtaca	1620
tttctaattg	cttgctactt	tttggtctga	caatgttaat	ccatattata	ggttgtaggc	1680
actatgaatg	tataaaaaaa	aaaaaaaaa				1708

<210> 2

<211> 1985

<212> DNA

<213> Canine MC4R Nucleotide Sequence

<400> 2

ctaagaccgt	ggggaggcag	ctgatgcgaa	catgtgcacg	cagattcagc	tcctgggtggc	60
tcggcgggcaa	ctcggagaat	tacttgcaac	agacctcact	gaatgcccta	gactaaagtt	120
aaggtggggag	tgaggacaaa	aaaaaaaaag	aaaaagaaaa	aagaaaaaaa	gaaaaaaaag	180
aaaaagcaaa	gagcagactc	tttgaactaa	gaatgagcat	ttcagaaatc	gaagatgtta	240
cagtgaaggt	gatcggagct	gtacctggaa	gacagtaaga	gctccactgc	cagccttttg	300
gagcacggga	caggtaactc	acacctggca	ggccagctgg	atcctcagaa	ctttgggacg	360
cacggagagg	gggagaacat	caccggggct	ccctggctgg	agaggccgaa	tcagtcccga	420
gggggtctgc	atacacttgt	tgcaggatga	actccacctt	tcagcacgga	atgcacactt	480
ctctccactt	ctggaaccgc	agcacctacg	gacagcacgg	caacgccact	gagtcctttg	540
gcaaaggcta	ccccgacggg	ggatgctacg	agcaactctt	cgtctccccg	gaggtgttcg	600
tgactctggg	ggtcataaag	ttgctggaga	acattctggt	gatcgtggca	atagccaaga	660
acaagaatct	gactcaccct	atgtactttt	tcatctgtag	cctggctgtg	gccgatatgc	720
tggtgagcgt	ttccaacggg	tcagagacca	tcgtcatcac	cctgttgaac	agtacggata	780
cggacgcgca	gagtttcacg	gtgaatattg	ataatgtcat	tgactcggtg	atctgtagct	840
ccttgctcgc	ctcgattttg	agcctgctct	caattgcagt	ggacaggtag	tttactatct	900
tttatgccct	ccagtaccat	aacatcatga	cggtgaggcg	ggttgggata	atcatcagtt	960
gcatctgggc	ggcttgcaag	gtgtcaggca	tcttgttcat	catttactcg	gacagtactg	1020
ctgtcatcat	ctgcctcatc	accatgttct	tcaccatgct	ggccctcatg	gcttctctct	1080
acgtccacat	gttcctcatg	gccagactgc	acatcaagag	aatcgccgtc	ctcccgggca	1140
cgggcaccat	ccgccaaggg	gccaacatga	agggtgccat	taccttgacc	atactcattg	1200
gggtcttcgt	cgtctgctgg	gctccattct	tcctccactt	gatattctac	atctcttgct	1260
cccagaatcc	atactgtgtg	tgcttcatgt	ctcactttta	cttgtaacct	attctgatca	1320
tgtgtaactc	catcatcgac	cctctcattt	atgcactccg	gagccaagag	ctgaggaaaa	1380
ccttcaaaga	gatcatctgt	tgctatcctc	tgggtggcct	ttgtgacttg	tctagcagat	1440
actagctggg	gacagaggaa	gtactaaaaa	catgcaccag	agacttcttc	atcctcacac	1500
aacatgaact	gtgtgcttgg	acaacagctg	cttcttcagt	ataaggcagg	agttgagaat	1560
atctgttgca	caaattcaac	tttatgatgt	tttgatgtga	aaaaaaaaat	gccagggtc	1620
tgtacattgc	taatgtcatg	ctacttttgg	gctgtgcatt	gttaatccat	ttcgacgctg	1680
tagacacttt	gaatttctag	aaaagaaaaa	agcttccatt	aaaagcatat	cagtgtttct	1740
tgttattcac	gaggatttgg	cactttgctt	gctttaggaa	acatagaaat	catagaatca	1800
ttaactatgt	agcctgataa	gtaacttctt	atattatact	atatcacatg	aaatgtgcag	1860
atgtgaatgt	agcatggggg	gtggatattg	aacaatagat	acttggtcat	taaaacaatc	1920
aactgaaatt	ttaagtaata	aatgtgttcc	attctccctg	ttgcagaaat	aaaaaaaaaa	1980

aaaaa

1985

<210> 3  
<211> 332  
<212> PRT  
<213> Feline MC4R protein Sequence

<400> 3  
Met Asn Ser Thr His His His Gly Met His Thr Ser Leu His Phe Trp  
1 5 10 15

Asn Arg Ser Thr Tyr Gly Pro His Ser Asn Ala Ser Glu Ser Leu Gly  
20 25 30

Lys Gly Tyr Ser Asp Gly Gly Cys Tyr Glu Gln Leu Phe Val Ser Pro  
35 40 45

Glu Val Phe Val Thr Leu Gly Val Ile Ser Leu Leu Glu Asn Ile Leu  
50 55 60

Val Ile Val Ala Ile Ala Lys Asn Lys Asn Leu His Ser Pro Met Tyr  
65 70 75 80

Phe Phe Ile Cys Ser Leu Ala Val Ala Asp Met Leu Val Ser Val Ser  
85 90 95

Asn Gly Ser Glu Thr Ile Val Ile Thr Leu Leu Asn Ser Thr Asp Thr  
100 105 110

Asp Ala Gln Ser Phe Thr Val Asn Ile Asp Asn Val Ile Asp Ser Val  
115 120 125

Ile Cys Ser Ser Leu Leu Ala Ser Ile Cys Ser Leu Leu Ser Ile Ala  
130 135 140

Val Asp Arg Tyr Phe Thr Ile Phe Tyr Ala Leu Gln Tyr His Asn Ile  
145 150 155 160

Met Thr Val Arg Arg Val Gly Ile Ile Ile Ser Cys Ile Trp Ala Ala  
165 170 175

Cys Thr Val Ser Gly Val Leu Phe Ile Ile Tyr Ser Asp Ser Ser Ala  
180 185 190

Val Ile Ile Cys Leu Ile Thr Met Phe Phe Thr Met Leu Ala Leu Met  
195 200 205



Asn Gly Ser Glu Thr Ile Val Ile Thr Leu Leu Asn Ser Thr Asp Thr  
 100 105 110

Asp Ala Gln Ser Phe Thr Val Asn Ile Asp Asn Val Ile Asp Ser Val  
 115 120 125

Ile Cys Ser Ser Leu Leu Ala Ser Ile Cys Ser Leu Leu Ser Ile Ala  
 130 135 140

Val Asp Arg Tyr Phe Thr Ile Phe Tyr Ala Leu Gln Tyr His Asn Ile  
 145 150 155 160

Met Thr Val Arg Arg Val Gly Ile Ile Ile Ser Cys Ile Trp Ala Ala  
 165 170 175

Cys Thr Val Ser Gly Ile Leu Phe Ile Ile Tyr Ser Asp Ser Thr Ala  
 180 185 190

Val Ile Ile Cys Leu Ile Thr Met Phe Phe Thr Met Leu Ala Leu Met  
 195 200 205

Ala Ser Leu Tyr Val His Met Phe Leu Met Ala Arg Leu His Ile Lys  
 210 215 220

Arg Ile Ala Val Leu Pro Gly Thr Gly Thr Ile Arg Gln Gly Ala Asn  
 225 230 235 240

Met Lys Gly Ala Ile Thr Leu Thr Ile Leu Ile Gly Val Phe Val Val  
 245 250 255

Cys Trp Ala Pro Phe Phe Leu His Leu Ile Phe Tyr Ile Ser Cys Pro  
 260 265 270

Gln Asn Pro Tyr Cys Val Cys Phe Met Ser His Phe Asn Leu Tyr Leu  
 275 280 285

Ile Leu Ile Met Cys Asn Ser Ile Ile Asp Pro Leu Ile Tyr Ala Leu  
 290 295 300

Arg Ser Gln Glu Leu Arg Lys Thr Phe Lys Glu Ile Ile Cys Cys Tyr  
 305 310 315 320

Pro Leu Gly Gly Leu Cys Asp Leu Ser Ser Arg Tyr  
 325 330

<210> 5

<211> 6148

<212> DNA

<213> pcDNA3.1zeo/murine G-alpha 15

<400> 5

```
gacggatcgg gagatctccc gatcccctat ggtcgactct cagtacaatc tgctctgatg 60
ccgcatagtt aagccagtat ctgctccctg cttgtgtgtt ggaggtcgct gagtagtgcg 120
cgagcaaaat ttaagctaca acaaggcaag gcttgaccga caattgcatg aagaatctgc 180
ttagggttag gcgtttttgcg ctgcttcgcg atgtacgggc cagatatacg cggtgacatt 240
gattattgac tagttattaa tagtaatcaa ttacgggggc attagttcat agcccatata 300
tggagttccg cgttacataa cttacggtaa atggcccgcg tggctgaccg cccaacgacc 360
ceegcccatt gaegtcaata atgaegtatg ttcccatagt aacgccaata gggactttcc 420
attgacgtca atgggtggac tatttacggg aaactgccc cttggcagta catcaagtgt 480
atcatatgcc aagtacgccc cctattgacg tcaatgacgg taaatggccc gcctggcatt 540
atgcccagta catgacctta tgggactttc ctacttgga gtacatctac gtattagtca 600
tcgctattac catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg 660
actcacgggg atttccaagt ctccacccca ttgacgtcaa tgggagtttg ttttggcacc 720
aaaatcaacg ggactttcca aaatgtcgta acaactccgc cccattgacg caaatgggcg 780
gtaggcgtgt acggtgggag gtctatataa gcagagctct ctggctaact agagaacca 840
ctgcttactg gcttatcgaa attaatacga ctactatag ggagacccaa gctggctagc 900
gtttaaactt aagcttggtg gtctgtgaag cgcccaccat ggcccgtcc ctgacttggg 960
gctgctgtcc ctggtgcctg acagaggagg agaagactgc cgccagaatc gaccaggaga 1020
tcaacaggat tttgttgga cagaaaaaac aagagcgcg ggaattgaaa ctctgctgt 1080
tggggcctgg tgagagcggg aagagtacgt tcatcaagca gatgcgcac attcacggtg 1140
tgggctactc ggaggaggac cgcagagcct tccggtgct catctaccag aacatcttcg 1200
tctccatgca ggccatgata gatgcgatgg accggtgca gatccccctc agcaggcctg 1260
acagcaagca gcacgccagc ctagtgatga ccaggaccc ctataaagtg agcacattcg 1320
agaagccata tgcagtggcc atgcagtacc tgtggcgga cgcgggcatc cgtgcatgct 1380
acgagcgaag gcgtgaattc caccttctgg actccgcggt gtattacctg tcacacctg 1440
agcgcataatc agaggacagc tacatcccca ctgcgcaaga cgtgctgcgc agtcgcatgc 1500
ccaccacagg catcaatgag tactgttct ccgtgaagaa aaccaaactg cgcacgtggtg 1560
atgttggtgg ccagaggtca gagcgtagga aatggattca ctgtttcgag aacgtgattg 1620
ccctcatcta cctggcctcc ctgagcgagt atgaccagt cctagaggag aacgatcagg 1680
agaaccgcat ggaggagagt ctgcgtctgt tcagcacgat cctagagctg ccctggttca 1740
agagcaacctc ggtcatcctc ttctcaaca agacggacat cctggaagat aagattcaca 1800
cctccacact ggccacatac ttcccagct tccagggacc ccggcgagac gcagaggccg 1860
ccaagagctt catcttgga atgtatgcgc gcgtgtacgc gagctgcgca gagccccagg 1920
acggtggcag gaaaggctcc cgcgcgcgcc gcttcttcgc acacttcacc tgtgccacgg 1980
acacgcaaag cgtccgcagc gtgttcaagg acgtgcggga ctcggtgctg gcccggtacc 2040
tggacgagat caacctgctg tgacgcagat ctaaagccga attctgcaga tatccatcac 2100
actggcgggc gctcgagcat gcatctagag ggcccgttta aaccgctga tcagcctcga 2160
ctgtgccttc tagttgccag ccactctgtt tttgccctc ccccgctgc tcttgaccc 2220
tggaaggtgc cactcccact gtcctttcct aataaaatga ggaaattgca tcgcattgtc 2280
tgagtaggtg tcattctatt ctggggggtg ggggtgggca ggacagcaag ggggaggatt 2340
gggaagacaa tagcaggcat gctggggatg cgggtgggctc tatggcttct gaggcgaaa 2400
gaaccagctg gggctctagg gggatatccc acgcgcctg tagcggcgca ttaagcgcg 2460
cgggtgtggt gggttacgcgc agcgtgaccg ctacacttgc cagcgcccta gcgcccgtc 2520
ctttcgcttt cttcccttcc tttctcgcca cgttcgccgg ctttccccgt caagctctaa 2580
```

atcggggcat	ccctttaggg	ttccgattta	gtgctttacg	gcacctcgac	cccaaaaaac	2640
ttgattaggg	tgatggttca	cgtagtgggc	catcgccctg	atagacggtt	tttcgcctt	2700
tgacgttgga	gtccacgttc	tttaatagt	gactcttggt	ccaaactgga	acaacactca	2760
accctatctc	ggtctattct	tttgatttat	aagggatttt	ggggatttcg	gcctattggt	2820
taaaaaatga	gctgatttaa	caaaaattta	acgcgaatta	attctgtgga	atgtgtgtca	2880
gttaggggtg	ggaaagtccc	caggctcccc	aggcaggcag	aagtatgcaa	agcatgcatc	2940
tcaattagtc	agcaaccagg	tgtggaaagt	ccccaggctc	cccagcaggc	agaagtatgc	3000
aaagcatgca	tctcaattag	tcagcaacca	tagtcccggc	cctaactccg	cccatcccgc	3060
ccctaactcc	gcccagttcc	gcccattctc	cgccccatgg	ctgactaatt	ttttttattt	3120
atgcagaggc	cgaggccggc	tctgcctctg	agctattcca	gaagtagtga	ggaggctttt	3180
ttggaggcct	aggcttttgc	aaaaagctcc	cgggagcttg	tatatccatt	ttcggatctg	3240
atcagcacgt	gttgacaatt	aatcatcggc	atagtatatc	ggcatagtat	aatacgacaa	3300
ggtgaggaac	taaaccatgg	ccaagttgac	cagtgccgtt	ccggtgctca	ccgcgcgcga	3360
cgtcgccgga	gcggtcgagt	tctggaccga	ccggctcggg	ttctcccggg	acttcgtgga	3420
ggacgacttc	gccggtgtgg	tccgggacga	cgtgacctg	ttcatcagcg	cggtccagga	3480
ccagggtggtg	ccggacaaca	ccctggcctg	ggtgtgggtg	cgcggcctgg	acgagctgta	3540
cgccgagtgg	tcggaggtcg	tgtccacgaa	cttccgggac	gcctccgggc	cggccatgac	3600
cgagatcggc	gagcagccgt	gggggcccga	gttcgcctg	cgcgacccgg	ccggcaactg	3660
cgtgcacttc	gtggccgagg	agcaggactg	acacgtgcta	cgagatttcg	attccaccgc	3720
cgccttctat	gaaaggttgg	gcttcggaat	cgttttccgg	gacgccggct	ggatgatcct	3780
ccagcgcggg	gatctcatgc	tggagttctt	cgccaccccc	aacttgttta	ttgcagctta	3840
taatggttac	aaataaagca	atagcatcac	aaatttcaca	aataaagcat	ttttttcact	3900
gcattctagt	tgtggtttgt	ccaaactcat	caatgtatct	tatcatgtct	gtataccgtc	3960
gacctctagc	tagagcttgg	cgtaatcatg	gtcatagctg	tttcctgtgt	gaaattgtta	4020
tccgctcaca	attccacaca	acatacgagc	cggaagcata	aagtgtaaag	cctgggggtgc	4080
ctaagtagt	agctaactca	cattaattgc	gttgcgctca	ctgcccgctt	tccagtcggg	4140
aaacctgtcg	tgccagctgc	attaatgaat	cggccaacgc	gcggggagag	gcggtttgcg	4200
tattgggcgc	tcttcgcgtt	cctcgctcac	tgactcgctg	cgcctcggtcg	ttcggctgcg	4260
gcgagcggta	tcagctcact	caaaggcggt	aatacggtta	tccacagaat	caggggataa	4320
cgcaggaaa	aacatgtgag	caaaggcca	gcaaaaggcc	aggaaccgta	aaaaggccgc	4380
gttgctggcg	tttttccata	ggctccgccc	ccctgacgag	catcacaaaa	atcgacgctc	4440
aagtcagagg	tggcgaaacc	cgacaggact	ataaagatac	caggcgtttc	cccctggaag	4500
ctccctcgtg	cgctctcctg	ttccgacct	gccgcttacc	ggatacctgt	ccgcctttct	4560
cccttcggga	agcgtggcgc	tttctcaatg	ctcacgctgt	aggtatctca	gttcgggtgta	4620
ggtcgttcgc	tccaagctgg	gctgtgtgca	cgaaccccc	gttcagcccc	accgctgcgc	4680
cttatccggt	aactatcgtc	ttgagtccaa	cccggtaaga	cacgacttat	cgcactggc	4740
agcagccact	ggtaacagga	ttagcagagc	gaggtagtga	ggcggtgcta	cagagttctt	4800
gaagtgggtg	cctaactacg	gtacactag	aaggacagta	tttggtatct	gcgctctgct	4860
gaagccagtt	accttcggaa	aaagagttgg	tagctcttga	tccggcaaac	aaaccaccgc	4920
tggtagcgg	ggtttttttg	tttgcaagca	gcagattacg	cgcagaaaaa	aaggatctca	4980
agaagatcct	ttgatctttt	ctacggggtc	tgacgctcag	tggaacgaaa	actcacgtta	5040
agggattttg	gtcatgagat	tatcaaaaag	gatcttcacc	tagatccttt	taaattaa	5100
atgaagtttt	aaatcaatct	aaagtatata	tgagtaaact	tggtctgaca	gttaccaatg	5160
cttaatcagt	gaggcaccta	tctcagcgat	ctgtctat	cgttcatcca	tagttgcctg	5220
actccccgtc	gtgtagataa	ctacgatacg	ggagggctta	ccatctggcc	ccagtgcctgc	5280
aatgataccg	cgagaccac	gtcacccggc	tccagattta	tcagcaataa	accagccagc	5340
cgggaaggcc	gagcgcagaa	gtggctcctgc	aactttatcc	gcctccatcc	agtctattaa	5400
ttgttgccgg	gaagctagag	taagtagttc	gccagttaat	agtttgcgca	acgttggttcg	5460

```

cattgctaca ggcacgctg tgtcacgctc gtcgtttggt atggcttcat tcagctccgg 5520
ttcccaacga tcaaggcgag ttacatgata ccccatgttg tgcaaaaaag cggttagctc 5580
cttcgggtcct ccgatcgttg tcagaagtaa gttggccgca gtgttatcac tcatggttat 5640
ggcagcactg cataattctc ttactgtcat gccatccgta agatgctttt ctgtgactgg 5700
tgagtactca accaagtcac tctgagaata gtgtatgcgg cgaccgagtt gctcttgccc 5760
ggcgtcaata cgggataata ccgcgccaca tagcagaact ttaaaagtgc tcatcattgg 5820
aaaacgttct tcggggcgaa aactctcaag gatcttaccg ctgttgagat ccagttcgat 5880
gtaaccact cgtgcaccca actgatcttc agcatctttt actttcacca gcgtttctgg 5940
gtgagcaaaa acaggaaggc aaaatgccgc aaaaaagga ataaggcgca cacggaaatg 6000
ttgaatactc atactcttcc tttttcaata ttattgaagc atttatcagg gttattgtct 6060
catgagcgga tacatatattg aatgtattta gaaaaataaa caaatagggg ttccgcgcac 6120
atttccccga aaagtgccac ctgacgtc 6148

```

```

<210> 6
<211> 6149
<212> DNA
<213> pcDNA3.1zeo/human G-alpha 16

```

```

<400> 6
gacggatcgg gagatctccc gatcccctat ggtcgactct cagtacaatc tgctctgatg 60
ccgcatagtt aagccagtat ctgctccctg cttgtgtggt ggaggtcgct gagtagtgcg 120
cgagcaaaat ttaagctaca acaaggcaag gcttgaccga caattgcatg aagaatctgc 180
ttagggttag gcgttttgcg ctgcttcgcg atgtacgggc cagatatacg cgttgacatt 240
gattattgac tagttattaa tagtaatcaa ttacgggggc attagttcat agcccatata 300
tggagttccg cgttacataa cttacggtaa atggcccgcg tggctgaccg cccaacgacc 360
ccgcgccatt gacgtcaata atgacgtatg ttcccatagt aacgcccaata gggactttcc 420
attgacgtca atgggtggac tatttacggt aaactgccca cttggcagta catcaagtgt 480
atcatatgcc aagtacgccc cctattgacg tcaatgacgg taaatggccc gcctggcatt 540
atgccagta catgacctta tgggactttc ctacttgga gtacatctac gtattagtca 600
tcgctattac catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg 660
actcacgggg atttccaagt ctccacccca ttgacgtcaa tgggagtttg ttttggcacc 720
aaaatcaacg ggactttcca aaatgtcgta acaactccgc cccattgacg caaatgggcg 780
gtaggcgtgt acggtgggag gtctatataa gcagagctct ctggctaact agagaacca 840
ctgcttactg gcttatcgaa attaatacga ctactatag ggagacccaa gctggctagc 900
gtttaaactt aagcttgact gaggccaccg caccatggcc cgctcgctga cctggcgctg 960
ctgcccctgg tgctgacgg aggatgagaa ggccgcgcgc cgggtggacc aggagatcaa 1020
caggatcctc ttggagcaga agaagcagga ccgcggggag ctgaagctgc tgcttttggg 1080
cccaggcgag agcgggaaga gcaccttcat caagcagatg cggatcatcc acggcgccgg 1140
ctactcgag gagagcgca agggcttcg gccctgggc taccagaaca tcttcgtgtc 1200
catgcggggc atgatcgagg ccatggagcg gctgcagatt ccattcagca ggcccgagag 1260
caagcaccac gctagcctgg tcatgagcca ggaccctat aaagtgacca cgtttgagaa 1320
gcgctacgct gcggccatgc agtggtgtg gagggatgcc ggcatccggg cctgctatga 1380
gcgtcgcgcg gaattccacc tgctcgattc agcgtgtac tacctgtccc acctggagcg 1440
catcaccgag gagggctacg tccccacagc tcaggacgtg ctccgcagcc gcatgccac 1500
cactggcatc aacgagtact gcttctccgt gcagaaaacc aacctgcgga tcgtggacgt 1560
cgggggcccag aagtcagagc gtaagaaatg gatccattgt ttcgagaacg tgatcgccct 1620
catctacctg gcctcactga gtgaatacga ccagtgcctg gaggagaaca accaggagaa 1680

```



ccgcatgaag	gagagcctcg	cattgttttg	gactatcctg	gaactaccct	ggttcaaaaag	1740
cacatccgtc	atcctctttc	tcaacaaaac	cgacatcctg	gaggagaaaa	tccccacctc	1800
ccacctgggt	acctatttcc	ccagtttcca	gggccctaag	caggatgctg	aggcagccaa	1860
gaggttcate	ctggacatgt	acacgaggat	gtacaccggg	tgcgtggacg	gccccgaggg	1920
cagcaagaag	ggcgacgat	cccagcgct	tttcagccac	tacacatgtg	ccacagacac	1980
acagaacatc	cgcaagggtc	tcaaggacgt	gcgggactcg	gtgctcgccc	gctacctgga	2040
cgagatcaac	ctgctgtgac	ccagatctaa	agccgaattc	tgcagatata	catcacactg	2100
gcggccgctc	gagcatgcat	ctagactaga	gggcccgttt	aaaccgcgtg	atcagcctcg	2160
actgtgcctt	ctagttgcca	gccatctgtt	gtttgcccct	cccccgtgcc	ttccttgacc	2220
ctggaagggt	ccactcccac	tgtcctttcc	taataaaaatg	aggaaattgc	atcgcatgtg	2280
ctgagtaggt	gtcattctat	tctggggggg	gggggtggggc	aggacagcaa	gggggaggat	2340
<del>tggaagaca</del>	<del>atagcaggca</del>	<del>tgctggggat</del>	<del>gcggtgggct</del>	<del>ctatggcttc</del>	<del>tgaggcgga</del>	<del>2400</del>
agaaccagct	ggggctctag	ggggtatccc	cacgcgccct	gtagcggcgc	attaagcgcg	2460
gcgggtgtgg	tggttacgcg	cagcgtgacc	gctacacttg	ccagcgccct	agcgcccgct	2520
cctttcgctt	tcttcccttc	ctttctcgcc	acgttcgccc	gctttccccg	tcaagctcta	2580
aatcggggca	tccctttagg	gttccgattt	agtgttttac	ggcacctcga	ccccaaaaaa	2640
cttgattagg	gtgatggttc	acgtagtggg	ccatcgccct	gatagacggg	ttttcgccct	2700
ttgacgttgg	agtccacggt	ctttaatagt	ggactcttgt	tccaaactgg	aacaacactc	2760
aaccctatct	cggctctattc	ttttgattta	taagggattt	tggggatttc	ggcctattgg	2820
ttaaaaaatg	agctgattta	acaaaaattt	aacgcgaatt	aattctgtgg	aatgtgtgtc	2880
agttaggggt	tggaaaagtc	ccaggtccc	caggcaggca	gaagtatgca	aagcatgcat	2940
ctcaattagt	cagcaaccag	gtgtggaaag	tccccaggct	ccccagcagg	cagaagtatg	3000
caaagcatgc	atctcaatta	gtcagcaacc	atagtcccgc	ccctaactcc	gcccattccc	3060
cccetaactc	cgcccagttc	cgcccattct	ccgcccctatg	gctgactaat	tttttttatt	3120
tatgcagagg	ccgaggccgc	ctctgectct	gagctattcc	agaagtagtg	aggaggcttt	3180
tttgaggagg	taggcttttg	caaaaagctc	ccgggagctt	gtatatccat	tttcggatct	3240
gatcagcacg	tgttgacaat	taatcatcgg	catagtatat	cggcatagta	taatacgaca	3300
aggtgaggaa	ctaaaccatg	gccaagttga	ccagtgcctg	tccgggtgctc	accgcgcgcg	3360
acgtcgccgg	agcggctcag	ttctggaccg	accggtcggg	gttctcccgg	gacttcgttg	3420
aggacgactt	cgccggtgtg	gtccgggacg	acgtgaccct	gttcattcagc	gcggtccagg	3480
accaggtggt	gcccggacaac	accctggcct	gggtgtgggt	gcgcggcctg	gacgagctgt	3540
acgcgagtg	gtcggaggtc	gtgtccacga	acttcgggga	cgccctccggg	ccggccatga	3600
ccgagatcgg	cgagcagccg	tgggggcggg	agttcgccct	gcgcgacccg	gccggcaact	3660
gcgtgcactt	cgtggccgag	gagcaggact	gacacgtgct	acgagatttc	gattccaccg	3720
ccgccttcta	tgaagaggtg	ggcttcggaa	tcgttttccg	ggacgcgggc	tggatgatcc	3780
tccagcgcgg	ggatctcatg	ctggagttct	tcgcccaccc	caacttgttt	attgcagctt	3840
ataatggtta	caaataaagc	aatagcatca	caaatttcac	aaataaagca	tttttttcac	3900
tgcattctag	ttgtggtttg	tccaaactca	tcaatgtatc	ttatcatgtc	tgtataccgt	3960
cgacctctag	ctagagcttg	gcgtaatcat	ggtcatagct	gtttcctgtg	tgaattgttt	4020
atccgctcac	aattccacac	aacatacgag	ccggaagcat	aaagtgtaaa	gctgggggtg	4080
cctaattgagt	gagctaaactc	acattaattg	cgttgcgctc	actgcccgct	ttccagtcgg	4140
gaaacctgtc	gtgccagctg	cattaatgaa	tcggccaacg	cgcggggaga	ggcggtttgc	4200
gtattggggc	ctcttccgct	tcctcgctca	ctgactcgct	gcgctcggtc	gttcggctgc	4260
ggcgagcggt	atcagctcac	tcaaaggcgg	taatacgggt	atccacagaa	tcaggggata	4320
acgcaggaaa	gaacatgtga	gcaaaaggcc	agcaaaaggc	caggaaccgt	aaaaaggccg	4380
cgttgctggc	gtttttccat	aggctccgcc	cccctgacga	gcatcacaaa	aatcgacgct	4440
caagtcagag	gtggcgaaac	ccgacaggac	tataaagata	ccaggcggtt	ccccctggaa	4500
gctccctcgt	gcgctctcct	gttccgaccc	tgcgcttac	cggatacctg	tccgcctttc	4560

6655-74666

tcccttcggg	aagcgtggcg	ctttctcaat	gctcacgctg	taggtatctc	agttcgggtgt	4620
aggtcgttcg	ctccaagctg	ggctgtgtgc	acgaaccccc	cgttcagccc	gaccgctgcg	4680
ccttatccgg	taactatcgt	cttgagtcca	acccggtaag	acacgactta	tcgccactgg	4740
cagcagccac	tggtaacagg	attagcagag	cgagggtatgt	aggcgggtgct	acagagttct	4800
tgaagtgggtg	gcctaactac	ggctacacta	gaaggacagt	atttggtatc	tgcgctctgc	4860
tgaagccagt	taccttcgga	aaaagagttg	gtagctcttg	atccggcaaa	caaaccaccg	4920
ctggtagcgg	tggttttttt	gtttgcaagc	agcagattac	gcgcaaaaaa	aaaggatctc	4980
aagaagatcc	tttgatcttt	tctacggggt	ctgacgctca	gtggaacgaa	aactcacggt	5040
aagggatttt	ggtcatgaga	ttatcaaaaa	ggatcttcac	ctagatcctt	ttaaattaaa	5100
aatgaagttt	taaatcaatc	taaagtatat	atgagtaaac	ttggtctgac	agttaccaat	5160
gcttaatcag	tgaggcacct	atctcagcga	tctgtctatt	tcgttcaccc	atagttgcct	5220
<del>gactccccgt</del>	<del>cgtgtagata</del>	<del>actacgatac</del>	<del>gggagggett</del>	<del>accatctggc</del>	<del>cccagtgcgt</del>	<del>5280</del>
caatgatacc	gcgagaccca	cgctcacccg	ctccagattt	atcagcaata	aaccagccag	5340
cgggaagggc	cgagcgcaga	agtggctctg	caactttatc	cgctccatc	cagtctatta	5400
attggtgccg	ggaagctaga	gtaagtagtt	cgccagttaa	tagtttgccg	aacgttggtg	5460
ccattgctac	aggcatcgtg	gtgtcacgct	cgctgtttgg	tatggcttca	ttcagctccg	5520
gttcccaacg	atcaaggcga	gttacatgat	ccccatggt	gtgcaaaaaa	gcggttagct	5580
ccttcgggtcc	tccgatcgtt	gtcagaagta	agttggccgc	agtgttatca	ctcatggtta	5640
tggcagcact	gcataattct	cttactgtca	tgccatccgt	aagatgcttt	tctgtgactg	5700
gtgagtactc	aaccaagtca	ttctgagaat	agtgtatgcg	gcgaccgagt	tgctcttgcc	5760
cggcgtcaat	acgggataat	accgcgccac	atagcagaac	tttaaaagtg	ctcatcattg	5820
gaaaacgttc	ttcggggcga	aaactctcaa	ggatcttacc	gctgttgaga	tccagttcga	5880
tgtaaccac	tcgtgcaccc	aactgatctt	cagcatcttt	tactttcacc	agcgtttctg	5940
ggtgagcaaa	aacaggaagg	caaatgccc	caaaaaagg	aataagggcg	acacggaaat	6000
gttgaatact	catactcttc	ctttttcaat	attattgaag	catttatcag	ggttattgtc	6060
tcatgagcgg	atacatattt	gaatgtattt	agaaaaataa	acaaataggg	gttccgcgca	6120
catttccccg	aaaagtgcga	cctgacgtc				6149